**MODULE 4: INTRODUCTION TO MULTIPLE LINEAR REGRESSION**

|  |  |
| --- | --- |
| 1. | Significance for the coefficients (b) is determined by   1. an F-test. 2. an *R*2 test. 3. a correlation coefficient. 4. a *t-*test. |
| 2. | Let a multiple linear model has a following equation:  Y = 1.4 + .00029 X1 + 2.4 X2 + 10.3 X3  Predicted value of Y:  If X1=3,000,000, X2=2, and X3=65, then   1. 1545.7 2. 1500 3. 1400 4. 1450 |
| 3. | Let a multiple linear model has a following equation:  Y = 1.4 + .00029 X1 + 2.4 X2 + 10.3 X3  a=1.4  This is the number of traffic fatalities that would be expected if all three independent variables (no population, no days snowed, and zero average speed) were equal to   1. One 2. Two 3. Constant 4. Zero |
| 4. | Let a multiple linear model has a following equation:  Y = 1.4 + .00029 X1 + 2.4 X2 + 10.3 X3  b1=.00029  If X2 and X3 remain the same, this indicates that for each extra person in the population, the number of yearly traffic fatalities .   1. Decreases by .00029 2. Increases by .00029 3. Decreases by 2.4 4. Increases by 2.4 |
| 5. | Let a multiple linear model has a following equation:  Y = 1.4 + .00029 X1 + 2.4 X2 + 10.3 X3  R2 = .78  We can explain --------------of the difference in annual fatality rates among states if we know the states' populations, days of snow, and average highway speeds.   1. 78% 2. 88% 3. 7.8% 4. .78% |
| 6. | A regression model in which more than one independent variable is used to predict the dependent variable is called\_\_\_   1. a simple linear regression model 2. a multiple regression model 3. an independent model 4. none of above |
| 7. | Significance for the coefficients (b) is determined by   1. F-test 2. R2 test 3. correlation coefficient 4. t-test |
| 8. | The R2 is the squared correlation of which two values?   1. y and the predicted values of y 2. y and each continuous x 3. b and t 4. b and se |
| 9. | Which of the following is true about the adjusted R2?   1. It is usually larger than the R2 2. It is only used when there is just one predictor 3. It is usually smaller than the R2 4. It is used to determine whether residuals are normally distributed |
| 10. | A multiple regression model has the form: y = 2 + 3x1 + 4x2. As x1 increases by 1 unit (holding x2 constant), y will\_\_\_\_   1. increase by 3 units 2. increase by 4 units 3. decrease by 3 units 4. decrease by 4 units |
| 11. | When there are more than one independent variables in the model, then the linear model is termed as \_\_\_\_\_\_\_   1. Unimodal 2. Multiple model 3. Multiple Linear model 4. Multiple Logistic model |
| 12. | The parameter β0 is termed as intercept term and the parameter β1 is termed as slope parameter. These parameters are usually called as \_\_\_\_\_\_\_\_\_   1. Regressionists 2. Coefficients 3. Regressive 4. Regression coefficients |
| 13. | \_\_\_\_\_\_\_\_ is a simple approach to supervised learning. It assumes that the dependence of Y on X1, X2, . . . Xp is linear.   1. Linear regression 2. Logistics regression 3. Gradient Descent 4. Greedy algorithms |
| 14. | What is predicting y for a value of x that is within the interval of points that we saw in the original data called?   1. Regression 2. Extrapolation 3. Intrapolation 4. Polation |
| 15. | Predicting y for a value of x that’s outside the range of values we actually saw for x in the original data is called   1. Regression 2. Extrapolation 3. Intrapolation 4. Polation |
| 16. | What does a multiple linear regression analysis examine?   1. The relationship between more than one dependent and only one independent variable 2. The relationship between one or more than one dependent and only one independent variable 3. The relationship between one dependent and more than one independent variables 4. The relationship between more than one independent variables |
| 17. | What does the following expression (H0: β1 = β2 = 0) mean?   1. One of the independent variables is useful in predicting the dependent variable 2. Both of the independent variables are useful in predicting the dependent variable 3. None of the independent variables is useful in predicting the dependent variable 4. There is a third independent variable predicting the dependent variable |
| 18. | Which of the following criteria is the most optimal for assessing the goodness of the fit of a multiple linear regression model?   1. Adjusted R2 2. R2 3. The intercept 4. The coefficient |
| 19. | A regression model in which more than one independent variable is used to predict the dependent variable is called   1. a simple linear regression model 2. a multiple regression model 3. an independent model 4. none of the above |
| 20. | A term used to describe the case when the independent variables in a multiple regression model are correlated is   1. regression 2. correlation 3. multicollinearity 4. none of the above |
| 21. | A multiple regression model has the form: y = 2 + 3x1 + 4x2. As x1 increases by 1 unit (holding x2 constant), y will   1. increase by 3 units 2. decrease by 3 units 3. increase by 4 units 4. decrease by 4 units |
| 22. | A multiple regression model has   1. only one independent variable 2. more than one dependent variable. 3. more than one independent variable 4. none of the above |
| 23. | The multiple coefficient of determination is computed by   1. dividing SSR by SST 2. dividing SSR by SST 3. none of the above 4. dividing SST by SSE |
| 24. | The ratio of MSR/MSE yields   1. the t statistic 2. the t statistic 3. the F statistic 4. none of the above |
| 25. | If the values of two variables move in the opposite direction, \_\_\_\_\_\_\_\_\_\_\_   1. The correlation is said to be linear 2. The correlation is said to be non-linear 3. The correlation is said to be positive 4. The correlation is said to be negative |
| 26. | Find zα/2 for 90% confidence level.   1. 1.63 2. 1.64 3. 1.65 4. 1.66 |
| 27. | Given *r*12 = 0.3, *r*13 = 0.5, *r*23 = 0.4 and σ1 = 3, σ2 = 4, σ3 = 5. Find the regression equation of *x*1 on *x*2 and *x*3 where *x*1, *x*2, and*x*3 have been measured from their actual means.   1. *x*1 = 0.9 *x*2 + 0.27 *x*3 2. *x*1 = 0.09 *x*2 + 0.27 *x*3 3. *x*1 = 0.09 *x*2 + 0.027 *x*3 4. *x*1 = 0.09 *x*2 + 2.7 *x*3 |
| 28. | If r12 = 0.28, r23 = 0.49, and r13 = 0.51, then the value of r12.3 is   1. 0.0401 2. 0.00401 3. 0.4001 4. 0.000401 |
| 29. | In MLR, the square of the multiple correlation coefficient or *R*2 is called the   1. Variance 2. Covariance 3. Cross-product 4. Coefficient of determination |
| 30. | The residual is defined as the difference between   1. the actual value of y and the estimated value of y 2. the actual value of x and the estimated value of x 3. the actual value of y and the estimated value of x 4. the actual value of x and the estimated value of y |
| 31. | Multiple linear regression (MLR) is a \_\_\_\_\_\_\_\_\_\_ type of statistical analysis   1. Univariate 2. Bivariate 3. Multivariate 4. Trivariate |
| 32. | When writing regression formulae, which of the following refers to the predicted value on the dependent variable (DV)?   1. Y 2. Y (hat) 3. X 4. X (hat) |
| 33. | Which of the following is true about the adjusted R2?   1. It is usually larger than the R2 2. It is only used when there is just one predictor 3. It is usually smaller than the R2 4. It is used to determine whether residuals are normally distributed |
| 34. | Least square method calculates the best-fitting line for the observed data by minimizing the sum of the squares of the \_\_\_\_\_\_\_ deviations.   1. Vertical 2. Horizontal 3. Both of these 4. None of these |
| 35. | A residual is defined as   1. The difference between the actual Y values and the mean of Y. 2. The difference between the actual Y values and the predicted Y values. 3. The predicted value of Y for the average X value. 4. The square root of the slope. |
| 36. | The correct relationship between SST, SSR, and SSE is given by  a) SSR = SST + SSE  b) SST = SSR + SSE  c) SSE = SSR – SST  d) all of the above |
| 37. | Below you are given a summary of the output from a simple linear regression analysis from a sample of size 15, SSR=100, SST = 152. The coefficient of determination is   1. 0.5200 2. 0.6579 3. 0.8111 4. 1.52 |

**MODULE 5: STATISTICAL INFERENCE**

|  |  |
| --- | --- |
| 1. | Which of the following is not property of Point Estimator   1. Bias 2. Consistency 3. Most efficient or unbiased 4. Asymmetrical distribution |
| 2. | Point Estimation, in [statistics](https://www.britannica.com/science/statistics), the process of finding an approximate -----------  of some parameter   1. Value 2. Integer 3. Population 4. Sample |
| 3. | A point estimator is defined as \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   1. Average of all sample values 2. Average of all population values 3. A single value that is best estimate of unknown population parameter 4. A single value from sample |
| 4. | Point of estimator with smaller standard error is said to have greater   1. Unbiasness 2. Biasness 3. Consistency 4. Efficiency |
| 5. | If expected value of sample statistics is equal to population parameter being estimated , sample statistics is said to be   1. Unbiased estimator 2. Biased estimator 3. Consistent estimator 4. Inconsistent estimator |
| 6. | What is mean by bias in point of estimator? choose correct statement   1. The bias of a point estimator is defined as the difference between the actual value of the estimator and the value of the parameter being estimated 2. The bias of a point estimator is defined as the difference between the expected value of the estimator and the value of the parameter being estimated 3. The bias of a point estimator is defined as the average of sample value 4. The bias of a point estimator is defined as the average of population value |
| 7. | A point estimator is defined as   1. average of sample values 2. average of population values 3. both 1 and 2 4. a single value that best estimate of an unknown population parameter |
| 8. | Inferential Statistics is the   1. process of using population parameter to estimate the values of for sample statistics 2. process of using sample statistics to estimate population parameter 3. Process that eliminates the problem of sampling error 4. All of the options |
| 9. | The difference between the sample value expected and the estimates value of the parameter is called as?   1. bias 2. error 3. contradiction 4. difference |
| 10. | Criteria to check a point estimator to be good are   1. Unbiasedness 2. Efficiency 3. Consistency 4. All of the options |
| 11. | The consistency of an estimator can be checked by comparing   1. Standard Deviation 2. Mean 3. Variance 4. Mean Square |
| 12. | A formula or rule used for estimating the parameter is called   1. Estimation 2. Estimate 3. Estimator 4. Interval estimate |
| 13. | Estimates given in the form of confidence intervals are called   1. Confidence limits 2. Degrees of freedom 3. Point estimates 4. Interval estimates |
| 14. | If the population standard deviation σ is unknown, and the sample size is small (n≤30),  the confidence interval for the population mean μ is based on   1. The binomial distribution 2. The t-distribution 3. The normal distribution 4. The Hypergeometric distribution |
| 15. | By increasing the sample size, the precision of confidence interval is \_\_\_\_\_\_   1. Zero 2. Remains same 3. Increased 4. Decreased |
| 16. | The distance between an estimate and the estimated parameter is called   1. Bias 2. Standard Error 3. Error of estimation 4. Sampling error |
| 17. | The number of values that are free to vary after a certain restriction are applied to the  data is called   1. Number of samples 2. Degrees of freedom 3. Confidence coefficient 4. Number of parameters |
| 18. | Estimation is of two types:   1. One sided and two sided 2. Type I and type II 3. Point estimation and interval estimation 4. Biased and unbiased |
| 19. | The process of using sample data to estimate the values of unknown population  parameter is called   1. Estimate 2. Estimator 3. Estimation 4. Interval estimation |
| 20. | The process of making estimates about the population parameter from a sample is  called   1. Statistical independence 2. Statistical inference 3. Statistical hypothesis 4. Statistical decision |
| 21. | Statistical inference has two branches namely   1. Level of confidence and degrees of freedom 2. Biased estimator and unbiased estimator 3. Point estimator and unbiased estimator 4. Estimation of parameter and testing of hypothesis |
| 22. | If the population standard deviation σ is known, the confidence interval for the  population mean μ is based on   1. The Poisson distribution 2. The t-distribution 3. The X2 -distribution 4. The normal distribution |
| 23. | Which of the following is not a property of a good estimator?   1. Biasedness 2. Sufficient 3. Efficient 4. Consistent |
| 24. | The standard normal distribution has   1. =1, = 0 2. = 0, = 1 3. = 0 , = 0 4. =1, = 1 |
| 25. | The distance between an estimate and the estimated parameter is called   1. Standard Error 2. Error of estimation 3. Sampling error 4. Bias |
| 26. | The maximum likelihood estimate is   1. minimum of α not necessarily in the parameter space 2. maximum of α in the parameter space 3. maximum of α not necessarily in the parameter space 4. minimum of α in the parameter space |
| 27. | \_\_\_\_\_\_\_\_\_\_\_ can be defined as a set of procedures based upon properly drawn samples to estimate population parameters using respective sample statistics.   1. Statistical Inferences 2. Estimation 3. Hypothesis Testing 4. Statistics |

**MODULE 6: TESTS OF HYPOTHESES**

|  |  |
| --- | --- |
| 1. | A deserving player is not selected in national team, it is an example of   1. Type-II error 2. Type-I error 3. Correct decision 4. Sampling error |
| 2. | The probability of rejecting a true hypothesis is called   1. Critical region 2. Level of significance 3. Test statistics 4. Statement of hypothesis |
| 3. | The region of acceptance of H0 is called   1. Critical region 2. Test statistics 3. Type-I error 4. Acceptance region |
| 4. | The probability of rejecting a false H0 is   1. Level of significance 2. Level of confidence 3. Critical region 4. Power of a test |
| 5. | Which statements is true if null hypothesis is false   1. Null hypothesis is accepted 2. Alternative hypothesis is accepted 3. Positive hypothesis is accepted 4. Negative hypothesis is accepted |
| 6. | The rejection probability of null hypothesis when it is true is called as?   1. Level of Confidence 2. Level of Significance 3. Level of Margin 4. Level of Rejection |
| 7. | The point where the null hypothesis gets rejected is called as?   1. Significant Value 2. Rejection Value 3. Acceptance Value 4. Critical Value |
| 8. | A statement about a population developed for the purpose of testing is called   1. hypothesis 2. hypothesis testing 3. test statistic 4. level of significance |
| 9. | A statement that is accepted if the sample data provide sufficient evidence that the null hypothesis is false is called\_\_\_   1. simple hypothesis 2. composite hypothesis 3. statistical hypothesis 4. alternative hypothesis |
| 10. | The range of test statistic-t is   1. 0 to 1 2. 0 to ∞ 3. -∞ to +∞ 4. -1 to +1 |
| 11. | Type 1 error occurs when \_\_\_\_\_\_\_\_\_\_\_\_\_\_   1. We reject H0 if it is True 2. We reject H0 if it is False 3. We accept H0 if it is True 4. We accept H0 if it is False |
| 12. | Power of a test is related to   1. Type-I error 2. Type-II error 3. Both 1 & 2 4. Neither 1 & 2 |
| 13. | Type 1 error occurs when?   1. We reject H0 if it is True 2. We reject H0 if it is False 3. We accept H0 if it is True 4. We accept H0 if it is False |
| 14. | Type 1 error occurs when?   1. We reject H0 if it is True 2. We reject H0 if it is False 3. We accept H0 if it is True 4. We accept H0 if it is False |
| 15. | If the critical region is evenly distributed then the test is referred as?   1. Two tailed 2. One tailed 3. Three tailed 4. Zero tailed |
| 16. | Consider a hypothesis H0 where ϕ0 = 5 against H1 where ϕ1 > 5. The test is?   1. Right tailed 2. Left tailed 3. Center tailed 4. Cross tailed |
| 17. | Consider a hypothesis where H0 where ϕ0 = 23 against H1 where ϕ1 < 23. The test  is?   1. Right tailed 2. Left tailed 3. Center tailed 4. Cross tailed |
| 18. | The probability of Type 1 error is referred as?   1. 1-α 2. β 3. α 4. 1-β |
| 19. | The choice of one-tailed test and two-tailed test depends upon   1. Null hypothesis 2. Alternative hypothesis 3. Composite hypothesis 4. None of these |
| 20. | When critical region is located on both side of the curve, it is called   1. One tail test 2. Two-tailed test 3. Left tailed test 4. Right tailed test |
| 21. | The point where the null hypothesis gets rejected is called as?   1. Significant Value 2. Rejection Value 3. Acceptance Value 4. Critical Value |
| 22. | Null and alternative hypotheses are statements about   1. population parameters 2. sample parameters 3. sample statistics 4. it depends-sometimes population parameters and sometimes sample statistics |
| 23. | Convert a random variable X into a standard normal variate Z and calculate the probability of scores P(2.9 ≤ X ≤ 7.1) for a normal distribution with mean (μ) = 5 and standard deviation (σ) = 10.   1. 0.1764 2. 0.1864 3. 0.1964 4. 0.1664 |
| 24. | The case of testing a simple null hypothesis (H0: θ = θ0) against a composite alternative hypothesis (H1: θ ≠ θ0) for a predetermined α, and so, the best test for H0 is called the \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ of level α.   1. powerful test 2. most powerful test 3. uniformly most powerful test 4. uniformly powerful test |
| 25. | Whether a test is one sided or two sided depends on   1. Alternative hypothesis 2. Composite hypothesis 3. Null hypothesis 4. Simple hypothesis |
| 26. | Large sample theory is applicable when   1. n > 30 2. n < 30 3. n < 100 4. n < 1000 |
| 27. | Student’s ‘ t’ distribution was pioneered by   1. Karl Pearson 2. Laplace 3. R.A. Fisher 4. William S. Gosset |
| 28. | Mean is ------ point estimator of population respectively.   1. biasedness 2. unbiasedness 3. both a and b 4. Invalid property |
| 29. | For probability and MLE, what is unknown ………. respectively.   1. Data, Population parameter 2. Data, Data 3. Population Parameter, Population Parameter 4. Population parameter, Data |
| 30. | What are third and fourth order moments?   1. Mean, Variance 2. Median, Range 3. Skewkness, Kurtosis 4. Kurtosis, Standard Deviation |